

(19)日本国特許庁 (J P)

(12) 特 許 公 報 (B 2)

(11)特許番号

第2866086号

(45)発行日 平成11年(1999) 3 月 8 日

(24)登録日 平成10年(1998)12月18日

(51)Int.Cl.<sup>8</sup>

識別記号

F I

G 1 0 L 9/14  
7/06  
9/14

G 1 0 L 9/14  
7/06  
9/14

L  
  
G

発明の数 4 (全 5 頁)

(21)出願番号 特願昭60-236930

(22)出願日 昭和60年(1985)10月23日

(65)公開番号 特開昭62-95600

(43)公開日 昭和62年(1987) 5 月 2 日

審査請求日 平成 4 年(1992) 9 月24日

前置審査

(73)特許権者 999999999

株式会社リコー

東京都大田区中馬込 1 丁目 3 番 6 号

(72)発明者 北川 博雄

東京都大田区中馬込 1 丁目 3 番 6 号 株  
式会社リコー内

(72)発明者 佐々部 昭一

東京都大田区中馬込 1 丁目 3 番 6 号 株  
式会社リコー内

(74)代理人 弁理士 高野 明近

審査官 山下 剛史

最終頁に続く

(54)【発明の名称】 残差駆動型音声合成方式

1

(57)【特許請求の範囲】

1. 音声の生成モデルに基づき、スペクトル情報と音源情報とから、ピッチ周期を持たない無声音を合成する音声分析合成方法において、合成器の逆フィルタから抽出される残差波形を入力する残差波形入力手段と、無声音部の定常部の残差波形から一定時間長の波形を切り出す区間切出し手段と、前記無声音部の残差波形から切り出された一定時間長の波形をフーリエ変換、ランダム位相化、逆フーリエ変換することによって得られる時間波形を伸縮処理して音源波形を出力する音源波形出力手段とから成ることを特徴とする残差駆動型音声合成方法。

2. 音声の生成モデルに基づき、スペクトル情報と音源情報とから、ピッチ周期を持たない無声音を合成する音声分析合成方法において、合成器の逆フィルタから抽出される残差波形を入力する残差波形入力手段と、無声音

2

部の定常部の残差波形から一定時間長の波形を切り出す区間切出し手段と、該区間切出し手段により切り出された残差波形を正規化する正規化手段と、該正規化手段により正規化された波形を伸縮処理して音源波形を出力する音源波形出力手段とから成り、前記一定時間長の残差波形をランダムな位置、長さで切り出して接続することによって前記伸縮処理を行うことを特徴とする残差駆動型音声合成方法。

3. 音声の生成モデルに基づき、スペクトル情報と音源情報とから、ピッチ周期を持たない無声音を合成する音声分析合成方法において、合成器の逆フィルタから抽出される残差波形を入力する残差波形入力手段と、無声音部の定常部の残差波形から一定時間長の波形を切り出す区間切出し手段と、該区間切出し手段により切り出された残差波形を正規化する正規化手段と、該正規化手段に

10

## 3

より正規化された波形を伸縮処理して音源波形を出力する音源波形出力手段とから成り、前記一定時間長の残差波形を任意の位置で折り返して接続することによって伸縮処理を行うことを特徴とする残差駆動型音声合成方法。

4. 音声の生成モデルに基づき、スペクトル情報と音源情報とから、ピッチ周期を持たない無声音を合成する音声分析合成方法において、合成器の逆フィルタから抽出される残差波形を入力する残差波形入力手段と、無声音部の定常部の残差波形から一定時間長の波形を切り出す区間切出し手段と、該区間切出し手段により切り出された残差波形を正規化する正規化手段と、該正規化手段により正規化された波形を伸縮処理して音源波形を出力する音源波形出力手段とから成り、前記無声音部の残差波形から切り出された一定時間長の波形のスペクトルをランダム位相化した波形を数パターン作成し、該パターンを接続して十分長い波形を生成し、該波形の任意の位置から必要とされる長さを切り出して駆動音源とすることを特徴とする残差駆動型音声合成方法。

## 【発明の詳細な説明】

## 技術分野

本発明は、残差駆動による音声合成方式、より詳細には、高品質の合成音声を得るための駆動波形生成方式に関する。

## 従来技術

音声の分析合成方式においては、その生成モデルに基づき、スペクトルパラメータ（LPC、PARCOR、LSPなど）と音源信号を音声合成フィルタに与えて合成音声を得ている。スペクトルパラメータは、声道特性を全極モデルで近似することによって得られるものであり、このモデルに立脚すれば、残差波形のスペクトルは白色化される。従つて、有声音部の音源信号にはインパルス列、無声音部の音源信号には白色雑音が多く用いられてきた。しかしながら、実際の音声生成過程は、全極モデルとは一致しないため、残差波形のスペクトルは、完全に白色化されることはなく、無声音部の音源として白色雑音を用いた場合、十分な音質が得られないという欠点がある。

## 目的

本発明は、上述のごとき実情に鑑みてなされたもので、特に、無声音部の音源に残差波形を用いることによって、高品質な合成音声を得ることを目的としてなされたものである。

## 構成

本発明は、上記目的を達成するために、（１）、音声の生成モデルに基づき、スペクトル情報と音源情報とから、ピッチ周期を持たない無声音を合成する音声分析合成方法において、合成器の逆フィルタから抽出される残差波形を入力する残差波形入力手段と、無声音部の定常部の残差波形から一定時間長の波形を切り出す区間切出し手段と、前記無声音部の残差波形から切り出された一

## 4

定時間長の波形をフーリエ変換、ランダム位相化、逆フーリエ変換することによって得られる時間波形を伸縮処理して音源波形を出力する音源波形出力手段とから成ること、又は、（２）、音声の生成モデルに基づき、スペクトル情報と音源情報とから、ピッチ周期を持たない無声音を合成する音声分析合成方法において、合成器の逆フィルタから抽出される残差波形を入力する残差波形入力手段と、無声音部の定常部の残差波形から一定時間長の波形を切り出す区間切出し手段と、該区間切出し手段により切り出された残差波形を正規化する正規化手段と、該正規化手段により正規化された波形を伸縮処理して音源波形を出力する音源波形出力手段とから成り、前記一定時間長の残差波形をランダムな位置、長さで切り出して接続することによって前記伸縮処理を行うこと、又は、（３）、音声の生成モデルに基づき、スペクトル情報と音源情報とから、ピッチ周期を持たない無声音を合成する音声分析合成方法において、合成器の逆フィルタから抽出される残差波形を入力する残差波形入力手段と、無声音部の定常部の残差波形から一定時間長の波形を切り出す区間切出し手段と、該区間切出し手段により切り出された残差波形を正規化する正規化手段と、該正規化手段により正規化された波形を伸縮処理して音源波形を出力する音源波形出力手段とから成り、前記一定時間長の残差波形を任意の位置で折り返して接続することによって伸縮処理を行うこと、又は、（４）、音声の生成モデルに基づき、スペクトル情報と音源情報とから、ピッチ周期を持たない無声音を合成する音声分析合成方法において、合成器の逆フィルタから抽出される残差波形を入力する残差波形入力手段と、無声音部の定常部の残差波形から一定時間長の波形を切り出す区間切出し手段と、該区間切出し手段により切り出された残差波形を正規化する正規化手段と、該正規化手段により正規化された波形を伸縮処理して音源波形を出力する音源波形出力手段とから成り、前記無声音部の残差波形から切り出された一定時間長の波形のスペクトルをランダム位相化した波形を数パターン作成し、該パターンを接続して十分長い波形を生成し、該波形の任意の位置から必要とされる長さを切り出して駆動音源とすることを特徴としたものである。以下、本発明の実施例に基いて説明する。

第１図は、駆動音源作成用の一定時間長の音源波形を生成するための一実施例を説明するための図で、図中、１は残差波形入力部、２は区間切出し部、３は振幅正規化部、４は音源波形出力部で、無声音部の定常部から一定時間長の残差波形を切出し（波形Ａ）、その振幅を正規化して音源波形（波形Ｂ）とするものである。ただし、振幅正規化は、合成音声のパワー制御を容易にするために行うものであり、必ずしも必要ではない。

第２図は、他の実施例を示す図で、図中、５はFFT、６はランダム位相化部、７は逆FFT、その他第１図と同様の作用をする部分には、第１図の場合と同一の参照番

5

号が付してある。而して、この実施例は、区間切出しされた残差波形にフーリエ変換、ランダム位相化、逆フーリエ変換の処理を施して音源波形を生成するもので、スペクトルのランダム位相化によって、振幅の正規化も同時に行われる。

第3図は、一定時間長の音源波形に伸縮処理を施して駆動音源を生成するための実施例を説明するための図で、この実施例は、音源波形(A)、(B)、(C)の任意の位置から任意の長さの波形をいくつか切出し、これらを接続して目的とする時間長に一致する駆動音源(D)を生成するものである。この際、切出す波形の位置と長さのいずれか一方は、固定でもよいが、音源波形の同じ部分が長時間繰返されないようにすることが重要である。同一部分の繰返しは、波形のランダム性を損い、繰返し周期が合成音声に含まれてしまうことになる。なお、音源波形の接続部は、荷重平均等滑らかに波形をつなぐための処理を施してもよい。また、駆動音源のパワーは、合成音声のパワーに合わせて自由に制御してよい。

第4図は、他の実施例を説明するための図で、この実施例は、音源波形(A)～(C)の任意の位置で波形を折返して接続する操作を数回繰返し、目的とする時間長に一致する駆動音源(D)を生成するものである。この操作により、接続部における連続性が確保されることに

6

なる。なお、波形の折返しとは、時間軸の向きを反転させることに相当する。

第5図は、本発明の他の実施例を説明するための図で、図中、11は切出し残差部、12<sub>1</sub>～12<sub>3</sub>はランダム位相化部、13<sub>1</sub>～13<sub>3</sub>は音源波形部で、この実施例は、切出された残差波形に複数のランダム位相化処理を施し、これらの波形をあらかじめ接続しておくことによって、十分に長い音源波形(A)を作成しておき、この波形の任意の部分を切り出して駆動音源(B)を生成するものである。以上によつて、無声音部における残差スペクトルを保存した音声合成が達成できる。

効果

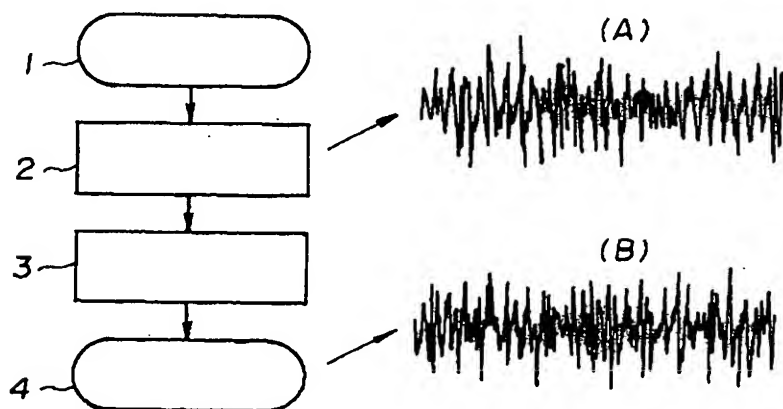
以上の説明から明らかなように、本発明によると、残差波形に含まれるスペクトル情報が駆動音源に保存されるため、白色雑音による駆動に比較して、より明瞭性、自然性に優れた高品質の合成音声を得られる。

【図面の簡単な説明】

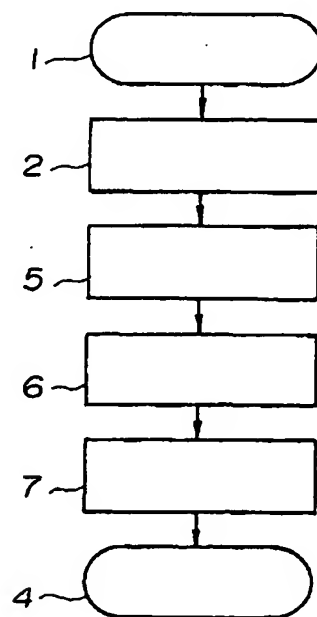
第1図乃至第5図は、それぞれ本発明の実施例を説明するための図である。

1……残差波形部、2……区間切出し部、3……振幅正規化部、4……音源波形部、5……FFT、6……ランダム位相化部、7……逆FFT、11……切出し残差部、12<sub>1</sub>～12<sub>3</sub>……ランダム位相化部、13<sub>1</sub>～13<sub>3</sub>……音源波形部。

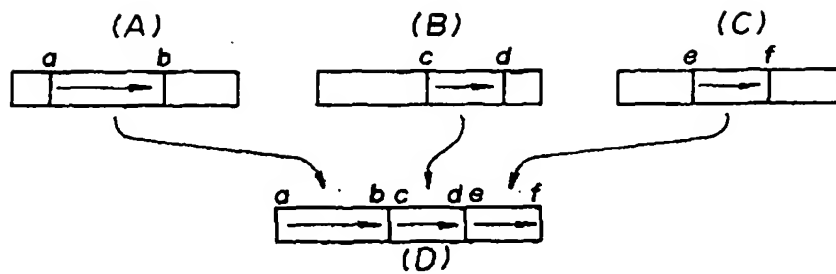
【第1図】



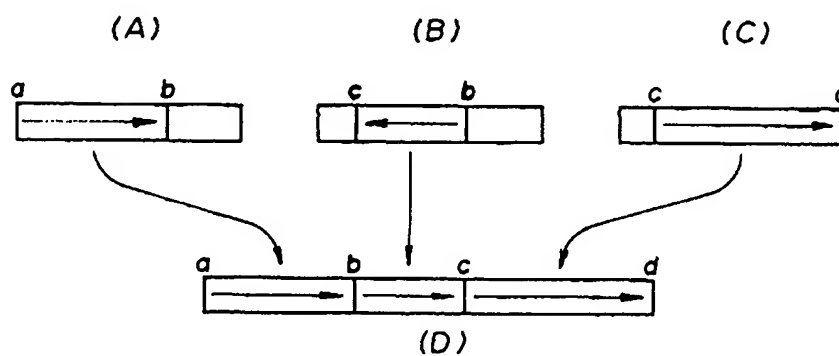
【第2図】



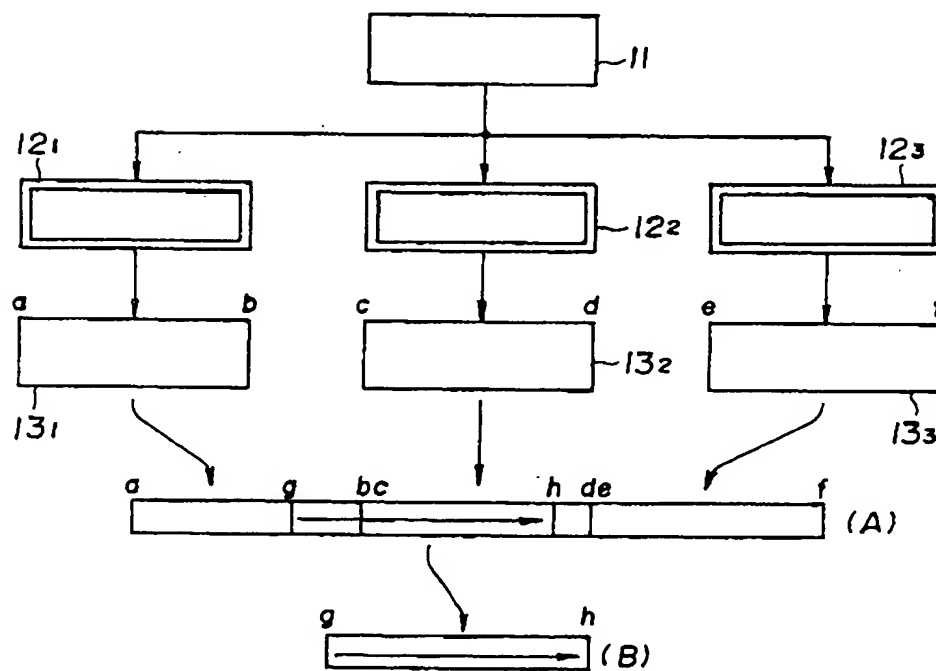
【第3図】



【第4図】



【第5図】



## フロントページの続き

(56) 参考文献 特開 昭58-154000 (J P, A)  
特開 昭59-172690 (J P, A)  
特開 昭59-61891 (J P, A)  
特開 昭58-88798 (J P, A)  
特開 昭57-104198 (J P, A)  
特公 昭55-36155 (J P, B 2)  
特公 昭59-5917 (J P, B 2)  
特公 昭59-14752 (J P, B 2)

(58) 調査した分野(Int. Cl.<sup>6</sup>, D B 名)  
G10L 7/00 - 7/06, 9/14